

(1) GENERAL INFORMATION:

- (i) APPLICANT: LUCAS, Sophie; BOON-FALLEUR, Thierry
- (ii) TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES CODING FOR TUMOR REJECTION ANTIGEN PRECURSORS OF MEMBERS OF THE MAGE-C AND MAGE-B FAMILIES AND USES THEREOF
- (iii) NUMBER OF SEQUENCES: 26
- (iv) CORRESPONDENCE ADDRESS:  
(A) ADDRESSEE: Fulbright & Jaworski L.L.P.  
(B) STREET: 801 Pennsylvania Avenue, N.W.  
(C) CITY: Washington  
(D) STATE: District of Columbia  
(E) COUNTRY: USA  
(F) ZIP: 20004
- (v) COMPUTER READABLE FORM:  
(A) MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
(B) COMPUTER: IBM PS/2  
(C) OPERATING SYSTEM: PC-DOS  
(D) SOFTWARE: Wordperfect
- (vi) CURRENT APPLICATION DATA:  
(A) APPLICATION NUMBER: US/09/501,104A  
(B) FILING DATE: 09-Feb-2000  
(C) CLASSIFICATION: 435
- (vii) PRIOR APPLICATION DATA:  
(A) APPLICATION NUMBER: 09/468,433  
(B) FILING DATE: December 17, 1999
- (vii) PRIOR APPLICATION DATA:  
(A) APPLICATION NUMBER: 09/066,281  
(B) FILING DATE: April 24, 1998
- (vii) PRIOR APPLICATION DATA:  
(A) APPLICATION NUMBER: 08/845,528  
(B) FILING DATE: April 25, 1997
- (viii) ATTORNEY/AGENT INFORMATION:  
(A) NAME: Mary Anne Schofield  
(B) REGISTRATION NUMBER: 36,669  
(C) REFERENCE/DOCKET NUMBER: LUD 5611.1 JEL/MAS
- (ix) TELECOMMUNICATION INFORMATION:  
(A) TELEPHONE: (212) 318-3100  
(B) TELEFAX: (212) 318-3400

10055403-201007

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4031 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double-stranded
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GGATCGTCTC AGGTCAGCGG AGGGAGGAGA CTTATAGACC TATCCAGTCT TCAAGGTGCT 60  
CCAGAAAGCA GGAGTTGAAG ACCTGGGTGT GAGGGACACA TACATCCTAA AAGCACCACA 120  
GCAGAGGAGG CCCAGGCAGT GCCAGGAGTC AAGGTTCCCA GAAGACAAAC CCCCTAGGAA 180  
GACAGGCGAC CTGTGAGGCC CTAGAGCACC ACCTTAAGAG AAGAAGAGCT GTAAGCCGGC 240  
CTTTGTCAGA GCCATCATGG GGGACAAGGA TATGCCTACT GCTGGGATGC CGAGTCTTCT 300  
CCAGAGTTCC TCTGAGAGTC CTCAGAGTTG TCCTGAGGGG GAGGACTCCC AGTCTCCTCT 360  
CCAGATTCCC CAGAGTTCTC CTGAGAGCGA CGACACCCTG TATCCTCTCC AGAGTCCTCA 420  
GAGTCGTTCT GAGGGGGAGG ACTCCTCGGA TCCTCTCCAG AGACCTCCTG AGGGGAAGGA 480  
CTCCCAGTCT CCTCTCCAGA TTCCCCAGAG TTCTCCTGAG GGCGACGACA CCCAGTCTCC 540  
TCTCCAGAAAT TCTCAGAGTT CTCCTGAGGG GAAGGACTCC CTGTCTCCTC TAGAGATTTC 600  
TCAGAGCCCT CCTGAGGGTG AGGATGTCCA GTCTCCTCTG CAGAATCCTG CGAGTTCCTT 660  
CTTCTCCTCT GCTTTATTGA GTATTTTCCA GAGTTCCCCT GAGAGAACTC AGAGTACTTT 720  
TGAGGGTTTT CCCCAGTCTC CTCTCCAGAT TCCTGTGAGC TCCTCCTCCT CCTCCACTTT 780  
ATTGAGTCTT TTCCAGAGTT CCCCTGAGAG AACTCAGAGT ACTTTTGAGG GTTTTCCCCA 840  
GTCTCTTCTC CAGATTCCTA TGACCTCCTC CTTCTCCTCT ACTTTATTGA GTATTTTCCA 900  
GAGTTCTCCT GAGAGTGCTC AAAGTACTTT TGAGGGTTTT CCCCAGTCTC CTCTCCAGAT 960  
TCCTGGGAGC CCCTCCTTCT CCTCCACTTT ACTGAGTCTT TTCCAGAGTT CCCCTGAGAG 1020  
AACTCACAGT ACTTTTGAGG GTTTTCCCCA GTCTCCTCTC CAGATTCCTA TGACCTCCTC 1080  
CTTCTCCTCT ACTTTATTGA GTATTTTCCA GAGTTCTCCT GAGAGTGCTC AAAGTACTTT 1140  
TGAGGGTTTT CCCCAGTCTC CTCTCCAGAT TCCTGGGAGC CCCTCCTTCT CCTCCACTTT 1200  
ACTGAGTCTT TTCCAGAGTT CCCCTGAGAG AACTCACAGT ACTTTTGAGG GTTTTCCCCA 1260  
GTCTCCTCTC CAGATTCCTA TGACCTCCTC CTTCTCCTCT ACTTTATTGA GTATTTTACA 1320  
GAGTTCTCCT GAGAGTGCTC AAAGTGCTTT TGAGGGTTTT CCCCAGTCTC CTCTCCAGAT 1380  
TCCTGTGAGC TCCTCTTTCT CCTACACTTT ATTGAGTCTT TTCCAGAGTT CCCCTGAGAG 1440  
AACTCAGAGT ACTTTTGAGG GTTTTCCCCA GTCTCCTCTC CAGATTCCTG TGAGCTCCTC 1500

100815001

CTCCTCCTCC TCCACTTTAT TGAGTCTTTT CCAGAGTTCC CCTGAGTGTA CTCAAAGTAC 1560  
TTTTGAGGGT TTTCCCCAGT CTCCTCTCCA GATTCCTCAG AGTCCTCCTG AAGGGGAGAA 1620  
TACCCATTCT CCTCTCCAGA TTGTTCCAAG TCTTCCTGAG TGGGAGGACT CCCTGTCTCC 1680  
TCACTACTTT CCTCAGAGCC CTCCTCAGGG GGAGGACTCC CTATCTCCTC ACTACTTTCC 1740  
TCAGAGCCCT CCTCAGGGGG AGGACTCCCT GTCTCCTCAC TACTTTCCTC AGAGCCCTCA 1800  
GGGGGAGGAC TCCCTGTCTC CTCACTACTT TCCTCAGAGC CCTCCTCAGG GGGAGGACTC 1860  
CATGTCTCCT CTCTACTTTC CTCAGAGTCC TCTTCAGGGG GAGGAATTCC AGTCTTCTCT 1920  
CCAGAGCCCT GTGAGCATCT GCTCCTCCTC CACTCCATCC AGTCTTCCCC AGAGTTTCCC 1980  
TGAGAGTTCT CAGAGTCCTC CTGAGGGGCC TGTCCAGTCT CCTCTCCATA GTCCTCAGAG 2040  
CCCTCCTGAG GGGATGCACT CCCAATCTCC TCTCCAGAGT CCTGAGAGTG CTCCTGAGGG 2100  
GGAGGATTCC CTGTCTCCTC TCCAAATTCC TCAGAGTCCT CTTGAGGGAG AGGACTCCCT 2160  
GTCTTCTCTC CATTTTCCTC AGAGTCCTCC TGAGTGGGAG GACTCCCTCT CTCCTCTCCA 2220  
CTTTCCTCAG TTTCTCCTC AGGGGGAGGA CTTCCAGTCT TCTCTCCAGA GTCCTGTGAG 2280  
TATCTGCTCC TCCTCCACTT CTTTGAGTCT TCCCCAGAGT TTCCCTGAGA GTCCTCAGAG 2340  
TCCTCCTGAG GGGCCTGCTC AGTCTCCTCT CCAGAGACCT GTCAGCTCCT TCTTCTCCTA 2400  
CACTTTAGCG AGTCTTCTCC AAAGTTCCCA TGAGAGTCCT CAGAGTCCTC CTGAGGGGCC 2460  
TGCCAGTCT CCTCTCCAGA GTCCTGTGAG CTCCTTCCCC TCCTCCACTT CATCGAGTCT 2520  
TTCCAGAGT TCTCCTGTGA GTCCTTCCC CTCCTCCACT TCATCGAGTC TTTCCAAGAG 2580  
TTCCCTGAG AGTCCTCTCC AGAGTCCTGT GATCTCCTTC TCCTCCTCCA CTTCATTGAG 2640  
CCCATTGAGT GAAGAGTCCA GCAGCCCAGT AGATGAATAT ACAAGTTCCT CAGACACCTT 2700  
GCTAGAGAGT GATTCCTTGA CAGACAGCGA GTCCTTGATA GAGAGCGAGC CCTTGTTTAC 2760  
TTATACACTG GATGAAAAGG TGGACGAGTT GCGCGGTTT CTTCTCCTCA AATATCAAGT 2820  
GAAGCAGCCT ATCACAAAGG CAGAGATGCT GACGAATGTC ATCAGCAGGT ACACGGGCTA 2880  
CTTTCCTGTG ATCTTCAGGA AAGCCCGTGA GTTCATAGAG ATACTTTTTG GCATTTCCCT 2940  
GAGAGAAGTG GACCCTGATG ACTCCTATGT CTTTGTAAC ACATTAGACC TCACCTCTGA 3000  
GGGGTGTCTG AGTGATGAGC AGGGCATGTC CCAGAACCGC CTCCTGATTC TTATTCTGAG 3060  
TATCATCTTC ATAAAGGGCA CCTATGCCTC TGAGGAGGTC ATCTGGGATG TGCTGAGTGG 3120  
AATAGGGGTG CGTGCTGGGA GGGAGCACTT TGCCTTTGGG GAGCCCAGGG AGCTCCTCAC 3180  
TAAAGTTTGG GTGCAGGAAC ATTACCTAGA GTACCGGGAG GTGCCCAACT CTTCTCCTCC 3240  
TCGTTACGAA TTCCTGTGGG GTCCAAGAGC TCATTGAGG GTCATTAAGA GGAAAGTAGT 3300

10035100.030102

AGAGTTTTTG GCCATGCTAA AGAATACCGT CCCTATTACC TTTCCATCCT CTTACAAGGA 3360  
TGCTTTGAAA GATGTGGAAG AGAGAGCCCA GGCCATAATT GACACCACAG ATGATTCGAC 3420  
TGCCACAGAA AGTGCAAGCT CCAGTGTCTAT GTCCCCCAGC TTCTCTTCTG AGTGAAGTCT 3480  
AGGGCAGATT CTTCCCTCTG AGTTTGAAGG GGGCAGTCGA GTTCTACGT GGTGGAGGGC 3540  
CTGGTTGAGG CTGGAGAGAA CACAGTGCTA TTTGCATTTT TGTTCCATAT GGGTAGTTAT 3600  
GGGGTTTACC TGTTTTACTT TTGGGTATTT TTCAAATGCT TTTCTATTA ATAACAGGTT 3660  
TAAATAGCTT CAGAATCCTA GTTTATGCAC ATGAGTCGCA CATGTATTGC TGTTTTTCTG 3720  
GTTTAAGAGT AACAGTTTGA TATTTTGTAA AAACAAAAC ACACCCAAAC ACACCACATT 3780  
GGGAAAACCT TCTGCCTCAT TTTGTGATGT GTCACAGGTT AATGTGGTGT TACTGTAGGA 3840  
ATTTTCTTGA AACTGTGAAG GAACTCTGCA GTTAAATAGT GGAATAAAGT AAAGGATTGT 3900  
TAATGTTTGC ATTTCTCAG GTCCTTTAGT CTGTTGTTCT TGAAACTAA AGATACATAC 3960  
CTGGTTTGCT TGGCTTACGT AAGAAAGTAG AAGAAAGTAA ACTGTAATAA ATAAAAAAAA 4020  
AAAAAAAAA A 4031

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single-stranded
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

GATCTGCGGT GA

12

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: SINGLE-stranded
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

GATCTGTTCA TG

12

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single-stranded
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

GATCTTCCCT CG

12

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 46 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single-stranded
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

NAACTGGAAG AATTCGCGGC CGCAGGAATT TTTTTTTTTT TTTTTT

46

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single-stranded
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (D) OTHER INFORMATION: BstX1 adapter upper strand

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

CTTTCAGCA CA

12

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1142
- (B) TYPE: amino acids
- (C) STRANDEDNESS: single-stranded
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

Met Gly Asp Lys Asp Met Pro Thr Ala Gly Met Pro Ser Leu Leu Gln  
5 10 15

Ser Ser Ser Glu Ser Pro Gln Ser Cys Pro Glu Gly Glu Asp Ser Gln  
20 25 30

Ser Pro Leu Gln Ile Pro Gln Ser Ser Pro Glu Ser Asp Asp Thr Leu  
35 40 45

20101010 "BOSTON" 10085400

10085103.030102

Tyr Pro Leu Gln Ser Pro Gln Ser Arg Ser Glu Gly Glu Asp Ser Ser  
50 55 60

Asp Pro Leu Gln Arg Pro Pro Glu Gly Lys Asp Ser Gln Ser Pro Leu  
65 70 75 80

Gln Ile Pro Gln Ser Ser Pro Glu Gly Asp Asp Thr Gln Ser Pro Leu  
85 90 95

Gln Asn Ser Gln Ser Ser Pro Glu Gly Lys Asp Ser Leu Ser Pro Leu  
100 105 110

Glu Ile Ser Gln Ser Pro Pro Glu Gly Glu Asp Val Gln Ser Pro Leu  
115 120 125

Gln Asn Pro Ala Ser Ser Phe Phe Ser Ser Ala Leu Leu Ser Ile Phe  
130 135 140

Gln Ser Ser Pro Glu Ser Ile Gln Ser Pro Phe Glu Gly Phe Pro Gln  
145 150 155 160

Ser Val Leu Gln Ile Pro Val Ser Ala Ala Ser Ser Ser Thr Leu Val  
165 170 175

Ser Ile Phe Gln Ser Ser Pro Glu Ser Thr Gln Ser Pro Phe Glu Gly  
180 185 190

Phe Pro Gln Ser Pro Leu Gln Ile Pro Val Ser Arg Ser Phe Ser Ser  
195 200 205

Thr Leu Leu Ser Ile Phe Gln Ser Ser Pro Glu Arg Ser Gln Arg Thr  
210 215 220

Ser Glu Gly Phe Ala Gln Ser Pro Leu Gln Ile Pro Val Ser Ser Ser  
225 230 235 240

Ser Ser Ser Thr Leu Leu Ser Leu Phe Gln Ser Ser Pro Glu Arg Thr  
245 250 255

Gln Ser Thr Phe Glu Gly Phe Pro Gln Ser Pro Leu Gln Ile Pro Val  
260 265 270

Ser Arg Ser Phe Ser Ser Thr Leu Leu Ser Ile Phe Gln Ser Ser Pro  
275 280 285

Glu Arg Thr Gln Ser Thr Phe Glu Gly Phe Ala Gln Ser Pro Leu Gln  
290 295 300

Ile Pro Val Ser Ser Ser Ser Ser Ser Thr Leu Leu Ser Leu Phe Gln  
305 310 315 320

Ser Ser Pro Glu Arg Thr Gln Ser Thr Phe Glu Gly Phe Pro Gln Ser  
325 330 335

Leu Leu Gln Ile Pro Met Thr Ser Ser Phe Ser Ser Thr Leu Leu Ser  
340 345 350

Ile Phe Gln Ser Ser Pro Glu Ser Ala Gln Ser Thr Phe Glu Gly Phe  
355 360 365







Ile Lys Gly Thr Tyr Ala Ser Glu Glu Val Ile Trp Asp Val Leu Ser  
1010 1015 1020

Gly Ile Gly Val Arg Ala Gly Arg Glu His Phe Ala Phe Gly Glu Pro  
1025 1030 1035 1040

Arg Glu Leu Leu Thr Lys Val Trp Val Gln Glu His Tyr Leu Glu Tyr  
1045 1050 1055

Arg Glu Val Pro Asn Ser Ser Pro Pro Arg Tyr Glu Phe Leu Trp Gly  
1060 1065 1070

Pro Arg Ala His Ser Glu Val Ile Lys Arg Lys Val Val Glu Phe Leu  
1075 1080 1085

Ala Met Leu Lys Asn Thr Val Pro Ile Thr Phe Pro Ser Ser Tyr Lys  
1090 1095 1100

Asp Ala Leu Lys Asp Val Glu Glu Arg Ala Gln Ala Ile Ile Asp Thr  
1105 1110 1115 1120

Thr Asp Asp Ser Thr Ala Thr Glu Ser Ala Ser Ser Ser Val Met Ser  
1125 1130 1135

Pro Ser Phe Ser Ser Glu  
1140

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1691 base pairs
- (B) TYPE: nucleotides
- (C) STRANDEDNESS: single stranded
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

CCATTCTGAG GGACGGCGTA GAGTTCGGCC GAAGGAACCT GACCCAGGCT CTGTGAGGAG 60

GCAAGGTTTT CAGGGGACAG GCCAACCCAG AGGACAGGAT TCCCTGGAGG CCACAGAGGA 120

GCACCAAGGA GAAGATCTGC CTGTGGGTCT TCATTGCCCCA GCTCCTGCCC ACACTCCTGC 180

CTGCTGCCCT GACGAGAGTC ATCATGTCTC TTGAGCAGAG GAGTCTGCAC TGCAAGCCTG 240

AGGAAGCCCT TGAGGCCCAA CAAGAGGCC TGGGCCTGGT GTGTGTGCAG GCTGCCACCT 300

CCTCCTCCTC TCCTCTGGTC CTGGGCACCC TGGAGGAGGT GCCCACTGCT GGGTCAACAG 360

ATCCTCCCCA GAGTCCTCAG GGAGCCTCCG CCTTTCCCAC TACCATCAAC TTCCTCGAC 420

AGAGGCAACC CAGTGAGGGT TCCAGCAGCC GTGAAGAGGA GGGGCCAAGC ACCTCTTGTA 480

TCCTGGAGTC CTTGTTCCGA GCAGTAATCA CTAAGAAGGT GGCTGATTG GTTGGTTTTT 540

TGCTCCTCAA ATATCGAGCC AGGGAGCCAG TCACAAAGGC AGAAATGCTG GAGAGTGTC 600

TCAAAAATTA CAAGCACTGT TTCCTGAGA TCTTCGGCAA AGCCTCTGAG TCCTTGACAG 660

TGGTCTTTGG CATTGACGTG AAGGAAGCAG ACCCCACCGG CCACTCCTAT GTCCTTGTCA 720  
 CCTGCCTAGG TCTCTCCTAT GATGGCCTGC TGGGTGATAA TCAGATCATG CCCAAGACAG 780  
 GCTTCCTGAT AATTGTCCTG GTCATGATTG CAATGGAGGG CGGCCATGCT CCTGAGGAGG 840  
 AAATCTGGGA GGAGCTGAGT GTGATGGAGG TGTATGATGG GAGGGAGCAC AGTGCCTATG 900  
 GGGAGCCCAG GAAGCTGCTC ACCCAAGATT TGGTGCAGGA AAAGTACCTG GAGTACCGGC 960  
 AGGTGCCGGA CAGTGATCCC GCACGCTATG AGTTCCTGTG GGGTCCAAGG GCCCTCGCTG 1020  
 AAACCAGCTA TGTGAAAGTC CTTGAGTATG TGATCAAGGT CAGTGCAAGA GTTCGCTTTT 1080  
 TCTTCCCATC CCTGCGTGAA GCAGCTTTGA GAGAGGAGGA AGAGGGAGTC TGAGCATGAG 1140  
 TTGCAGCCAA GGCCAGTGGG AGGGGGACTG GGCCAGTGCA CTTCCAGGG CCGCGTCCAG 1200  
 CAGCTTCCCC TGCCTCGTGT GACATGAGGC CCATTCTTCA CTCTGAAGAG AGCGGTCAGT 1260  
 GTTCTCAGTA GTAGGTTTCT GTTCTATTGG GTGACTTGGA GATTTATCTT TGTTCCTCTT 1320  
 TGGAATTGTT CAAATGTTTT TTTTAAAGG ATGGTTGAAT GAAC'TTCAGC ATCCAAGTTT 1380  
 ATGAATGACA GCAGTCACAC AGTTCTGTGT ATATAGTTTA AGGGTAAGAG TCTTGTGTTT 1440  
 TATTCAGATT GGGAAATCCA TTCTATTTTG TGAATTGGGA TAATAACAGC AGTGGGAATAA 1500  
 GTACTTAGAA ATGTGAAAAA TGAGCAGTAA AATAGATGAG ATAAAGAACT AAAGAAATTA 1560  
 AGAGATAGTC AATTCTTGCC TTATACCTCA GTCTATTCTG TAAAATTTTT AAAGATATAT 1620  
 GCATACCTGG ATTCCTTGG CTTCTTTGAG AATGTAAGAG AAATTAAATC TGAATAAAGA 1680  
 ATTCTTCCTG T 1691

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4225 base pairs
- (B) TYPE: nucleic acids
- (C) STRANDEDNESS: double-stranded
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

GGATCGTCTC AGGTCAGCGG AGGGAGGAGA CTTATAGACC TATCCAGTCT TCAAGGTGCT 60  
 CCAGAAAGCA GGAGTTGAAG ACCTGGGTGT GAGGGACACA TACATCCTAA AAGCACCACA 120  
 GCAGAGGAGG CCCAGGCAGT GCCAGGAGTC AAGGTTCCCA GAAGACAAAC CCCCTAGGAA 180  
 GACAGGCGAC CTGTGAGGCC CTAGAGCACC ACCTTAAGAG AAGAAGAGCT GTAAGCCGGC 240  
 CTTTGTGAGA GCCATCATGG GGGACAAGGA TATGCCTACT GCTGGGATGC CGAGTCTTCT 300  
 CCAGAGTTCC TCTGAGAGTC CTCAGAGTTG TCCTGAGGGG GAGGACTCCC AGTCTCCTCT 360  
 CCAGATTCCC CAGAGTTCTC CTGAGAGCGA CGACACCCTG TATCCTCTCC AGAGTCCTCA 420

GAGTCGTTCT	GAGGGGGAGG	ACTCCTCGGA	TCCTCTCCAG	AGACCTCCTG	AGGGGAAGGA	480
CTCCCAGTCT	CCTCTCCAGA	TTCCCCAGAG	TTCTCCTGAG	GGCGACGACA	CCCAGTCTCC	540
TCTCCAGAAT	TCTCAGAGTT	CTCCTGAGGG	GAAGGACTCC	CTGTCTCCTC	TAGAGATTTT	600
TCAGAGCCCT	CCTGAGGGTG	AGGATGTCCA	GTCTCCTCTG	CAGAATCCTG	CGAGTTCCTT	660
CTTCTCCTCT	GCTTTATTGA	GTATTTTCCA	GAGTTCCCCT	GAGAGTATTC	AAAGTCCTTT	720
TGAGGGTTTT	CCCCAGTCTG	TTCTCCAGAT	TCCTGTGAGC	GCCGCCTCCT	CCTCCACTTT	780
AGTGAGTATT	TTCCAGAGTT	CCCCTGAGAG	TACTCAAAGT	CCTTTTGAGG	GTTTTCCCCA	840
GTCTCCACTC	CAGATTCCTG	TGAGCCGCTC	CTTCTCCTCC	ACTTTATTGA	GTATTTTCCA	900
GAGTTCCCCT	GAGAGAAGTC	AGAGAACTTC	TGAGGGTTTT	GCACAGTCTC	CTCTCCAGAT	960
TCCTGTGAGC	TCCTCCTCGT	CCTCCACTTT	ACTGAGTCTT	TTCCAGAGTT	CCCCTGAGAG	1020
AACTCAGAGT	ACTTTTGAGG	GTTTTCCCCA	GTCTCCACTC	CAGATTCCTG	TGAGCCGCTC	1080
CTTCTCCTCC	ACTTTATTGA	GTATTTTCCA	GAGTTCCCCT	GAGAGAACTC	AGAGTACTTT	1140
TGAGGGTTTT	GCCCAGTCTC	CTCTCCAGAT	TCCTGTGAGC	TCCTCCTCCT	CCTCCACTTT	1200
ATTGAGTCTT	TTCCAGAGTT	CCCCTGAGAG	AACTCAGAGT	ACTTTTGAGG	GTTTTCCCCA	1260
GTCTCTTCTC	CAGATTCCTA	TGACCTCCTC	CTTCTCCTCT	ACTTTATTGA	GTATTTTCCA	1320
GAGTTCTCCT	GAGAGTGCTC	AAAGTACTTT	TGAGGGTTTT	CCCCAGTCTC	CTCTCCAGAT	1380
TCCTGGGAGC	CCCTCCTTCT	CCTCCACTTT	ACTGAGTCTT	TTCCAGAGTT	CCCCTGAGAG	1440
AACTCACAGT	ACTTTTGAGG	GTTTTCCCCA	GTCTCCTCTC	CAGATTCCTA	TGACCTCCTC	1500
CTTCTCCTCT	ACTTTATTGA	GTATTTTACA	GAGTTCTCCT	GAGAGTGCTC	AAAGTGCTTT	1560
TGAGGGTTTT	CCCCAGTCTC	CTCTCCAGAT	TCCTGTGAGC	TCCTCTTTCT	CCTACACTTT	1620
ATTGAGTCTT	TTCCAGAGTT	CCCCTGAGAG	AACTCAGAGT	ACTTTTGAGG	GTTTTCCCCA	1680
GTCTCCTCTC	CAGATTCCTG	TGAGCTCCTC	CTCCTCCTCC	TCCACTTTAT	TGAGTCTTTT	1740
CCAGAGTTCC	CCTGAGTGTA	CTCAAAGTAC	TTTTGAGGGT	TTCCCCCAGT	CTCCTCTCCA	1800
GATTCTCTCAG	AGTCCTCCTG	AAGGGGAGAA	TACCCATTCT	CCTCTCCAGA	TTGTTCCAAG	1860
TCTTCCTGAG	TGGGAGGACT	CCCTGTCTCC	TCACTACTTT	CCTCAGAGCC	CTCCTCAGGG	1920
GGAGGACTCC	CTATCTCCTC	ACTACTTTCC	TCAGAGCCCT	CCTCAGGGGG	AGGACTCCCT	1980
GTCTCCTCAC	TACTTTTCTC	AGAGCCCTCA	GGGGGAGGAC	TCCCTGTCTC	CTCACTACTT	2040
TCCTCAGAGC	CCTCCTCAGG	GGGAGGACTC	CATGTCTCCT	CTCTACTTTC	CTCAGAGTCC	2100
TCTTCAGGGG	GAGGAATTCC	AGTCTTCTCT	CCAGAGCCCT	GTGAGCATCT	GCTCCTCCTC	2160
CACTCCATCC	AGTCTTCCCC	AGAGTTTCCC	TGAGAGTTCT	CAGAGTCTTC	CTGAGGGGCC	2220

TGTCCAGTCT CCTCTCCATA GTCCTCAGAG CCCTCCTGAG GGGATGCACT CCCAATCTCC 2280  
TCTCCAGAGT CCTGAGAGTG CTCCTGAGGG GGAGGATTCC CTGTCTCCTC TCCAAATTCC 2340  
TCAGAGTCCT CTTGAGGGAG AGGACTCCCT GTCTTCTCTC CATTTTCCTC AGAGTCCTCC 2400  
TGAGTGGGAG GACTCCCTCT CTCCTCTCCA CTTTCCTCAG TTTCTCCTC AGGGGGAGGA 2460  
CTTCCAGTCT TCTCTCCAGA GTCCTGTGAG TATCTGCTCC TCCTCCACTT CTTTGAGTCT 2520  
TCCCCAGAGT TTCCCTGAGA GTCCTCAGAG TCCTCCTGAG GGGCCTGCTC AGTCTCCTCT 2580  
CCAGAGACCT GTCAGCTCCT TCTTCTCCTA CACTTTAGCG AGTCTTCTCC AAAGTTCCCA 2640  
TGAGAGTCCT CAGAGTCCTC CTGAGGGGCC TGCCAGTCT CCTCTCCAGA GTCCTGTGAG 2700  
CTCCTTCCCC TCCTCCACTT CATCGAGTCT TTCCAGAGT TCTCCTGTGA GCTCCTTCCC 2760  
CTCCTCCACT TCATCGAGTC TTTCCAAGAG TTCCCCTGAG AGTCCTCTCC AGAGTCCTGT 2820  
GATCTCCTTC TCCTCCTCCA CTTCAATTGAG CCCATTGAGT GAAGAGTCCA GCAGCCCAGT 2880  
AGATGAATAT ACAAGTTCCT CAGACACCTT GCTAGAGAGT GATTCTTGA CAGACAGCGA 2940  
GTCCTTGATA GAGAGCGAGC CCTTGTTTAC TTATACTG GATGAAAAGG TGGACGAGTT 3000  
GGCGCGGTTT CTTCTCCTCA AATATCAAGT GAAGCAGCCT ATCACAAAGG CAGAGATGCT 3060  
GACGAATGTC ATCAGCAGGT ACACGGGCTA CTTTCCTGTG ATCTTCAGGA AAGCCCGTGA 3120  
GTTTCATAGAG ATACTTTTTG GCATTTCCCT GAGAGAAGTG GACCCTGATG ACTCCTATGT 3180  
CTTTGTAAAC ACATTAGACC TCACCTCTGA GGGGTGTCTG AGTGATGAGC AGGGCATGTC 3240  
CCAGAACCGC CTCCTGATTC TTATTCTGAG TATCATCTTC ATAAAGGGCA CCTATGCCTC 3300  
TGAGGAGGTC ATCTGGGATG TGCTGAGTGG AATAGGGGTG CGTGCTGGGA GGGAGCACTT 3360  
TGCCTTTGGG GAGCCCAGGG AGCTCCTCAC TAAAGTTTGG GTGCAGGAAC ATTACCTAGA 3420  
GTACCGGGAG GTGCCCAACT CTTCTCCTCC TCGTTACGAA TTCCTGTGGG GTCCAAGAGC 3480  
TCATTCAGAA GTCATTAAGA GGAAAGTAGT AGAGTTTTTG GCCATGCTAA AGAATACCGT 3540  
CCCTATTACC TTTCCATCCT CTTACAAGGA TGCTTTGAAA GATGTGGAAG AGAGAGCCCA 3600  
GGCCATAATT GACACCACAG ATGATTCGAC TGCCACAGAA AGTGCAAGCT CCAGTGTCAT 3660  
GTCCCCCAGC TTCTCTTCTG AGTGAAGTCT AGGGCAGATT CTTCCCTCTG AGTTTGAAGG 3720  
GGGCAGTCGA GTTTCTACGT GGTGGAGGGC CTGGTTGAGG CTGGAGAGAA CACAGTGCTA 3780  
TTTGCATTTT TGTTCCATAT GGGTAGTTAT GGGGTTTACC TGTTTTACTT TTGGGTATTT 3840  
TTCAAATGCT TTTCTATTA ATAACAGGT TAAATAGCTT CAGAATCCTA GTTTATGCAC 3900  
ATGAGTCGCA CATGTATTGC TGTTTTTCTG GTTTAAGAGT AACAGTTTGA TATTTTGTA 3960  
AAACAAAAAC ACACCCAAAC ACACCACATT GGGAAAACCT TCTGCCTCAT TTTGTGATGT 4020

GTCACAGGTT AATGTGGTGT TACTGTAGGA ATTTTCTTGA AACTGTGAAG GAACTCTGCA 4080  
 GTTAAATAGT GGAATAAAGT AAAGGATTGT TAATGTTTGC ATTTCCCTCAG GTCCTTTAGT 4140  
 CTGTTGTTCT TGAAACTAA AGATACATAC CTGGTTTGCT TGGCTTACGT AAGAAAGTAG 4200  
 AAGAAAGTAA ACTGTAATAA ATAAA 4225

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 309
- (B) TYPE: amino acids
- (C) STRANDEDNESS: single stranded
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

Met Ser Leu Glu Gln Arg Ser Leu His Cys Lys Pro Glu Glu Ala Leu  
                   5                  10                  15  
 Glu Ala Gln Gln Glu Ala Leu Gly Leu Val Cys Val Gln Ala Ala Thr  
                   20                  25                  30  
 Ser Ser Ser Ser Pro Leu Val Leu Gly Thr Leu Glu Glu Val Pro Thr  
                   35                  40                  45  
 Ala Gly Ser Thr Asp Pro Pro Gln Ser Pro Gln Gly Ala Ser Ala Phe  
                   50                  55                  60  
 Pro Thr Thr Ile Asn Phe Thr Arg Gln Arg Gln Pro Ser Glu Gly Ser  
                   65                  70                  75                  80  
 Ser Ser Arg Glu Glu Gly Pro Ser Thr Ser Cys Ile Leu Glu Ser  
                   85                  90                  95  
 Leu Phe Arg Ala Val Ile Thr Lys Lys Val Ala Asp Leu Val Gly Phe  
                   100                  105                  110  
 Leu Leu Leu Lys Tyr Arg Ala Arg Glu Pro Val Thr Lys Ala Glu Met  
                   115                  120                  125  
 Leu Glu Ser Val Ile Lys Asn Tyr Lys His Cys Phe Pro Glu Ile Phe  
                   130                  135                  140  
 Gly Lys Ala Ser Glu Ser Leu Gln Leu Val Phe Gly Ile Asp Val Lys  
                   145                  150                  155                  160  
 Glu Ala Asp Pro Thr Gly His Ser Tyr Val Leu Val Thr Cys Leu Gly  
                   165                  170                  175  
 Leu Ser Tyr Asp Gly Leu Leu Gly Asp Asn Gln Ile Met Pro Lys Thr  
                   180                  185                  190  
 Gly Phe Leu Ile Ile Val Leu Val Met Ile Ala Met Glu Gly Gly His  
                   195                  200                  205

10085108.030102

Ala Pro Glu Glu Glu Ile Trp Glu Glu Leu Ser Val Met Glu Val Tyr  
 210 215 220

Asp Gly Arg Glu His Ser Ala Tyr Gly Glu Pro Arg Lys Leu Leu Thr  
 225 230 235 240

Gln Asp Leu Val Gln Glu Lys Tyr Leu Glu Tyr Arg Gln Val Pro Asp  
 245 250 255

Ser Asp Pro Ala Arg Tyr Glu Phe Leu Trp Gly Pro Arg Ala Leu Ala  
 260 265 270

Glu Thr Ser Tyr Val Lys Val Leu Glu Tyr Val Ile Lys Val Ser Ala  
 275 280 285

Arg Val Arg Phe Phe Phe Pro Ser Leu Arg Glu Ala Ala Leu Arg Glu  
 290 295 300

Glu Glu Glu Gly Val  
 305 309

(2) INFORMATION FOR SEQ ID NO: 11:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single-stranded
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

AGCACTCTCC AGCCTCTCAC CGCA

24

(2) INFORMATION FOR SEQ ID NO: 12:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single-stranded
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

ACCGACGTCG ACTATCCATG AACA

24

(2) INFORMATION FOR SEQ ID NO: 13:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single-stranded
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

20100101 01010101

AGGCAACTGT GCTATCCGAG GGAA

24

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single-stranded
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (D) OTHER INFORMATION: BstXI adapter lower strand

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

CTGGAAAG

8

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single-stranded
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

AGGCGCGAAT CAAGTTAG

18

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single-stranded
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

CTCCTCTGCT GTGCTGAC

18

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single-stranded
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

AGCTGCCTCT GGTGGCAGA

20

20100505 09:55:00

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1983 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double-stranded
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

TGGGAATCTG ACGGATCGGA GGCATTTGTG AGGAGGCGCG AATCAAGTTA GCGGGGGGAA 60  
GAGTCTTAGA CCTGGCCAGT CCTCAGGGTG AGGGCCCTGA GGAAGAACTG AGGGACCTCC 120  
CACCATAGAG AGAAGAAACC CCGGCCTGTA CTGCGCTGCC GTGAGACTGG TGCTCCAGGA 180  
ACCAGGTGGT GACGAACCTGG GTGTGAGGCA CACAGCCTAA AGTCAGCACA GCAGAGGAGG 240  
CCCAGGCAGT GCCAGGAGTC AAGGCCTGTT GGATCTCATC ATCCATATCC CTGTTGATAC 300  
GTTTACCTGC TGCTCCTGAA GAAGTCGTCA TGCCTCCCGT TCCAGGCGTT CCATTCCGCA 360  
ACGTTGACAA CGACTCCCCG ACCTCAGTTG AGTTAGAAGA CTGGGTAGAT GCACAGCATC 420  
CCACAGATGA GGAAGAGGAG GAAGCCTCCT CCGCCTCTTC CACTTTGTAC TTAGTATTTT 480  
CCCCCTCTTC TTTCTCCACA TCCTCTTCTC TGATTCTTGG TGGTCCTGAG GAGGAGGAGG 540  
TGCCCTCTGG TGTGATACCA AATCTTACCG AGAGCATTCC CAGTAGTCCT CCACAGGGTC 600  
CTCCACAGGG TCCTTCCCAG AGTCCTCTGA GCTCCTGCTG CTCCTCTTTT TCATGGAGCT 660  
CATTCAGTGA GGAGTCCAGC AGCCAGAAAG GGGAGGATAC AGGCACCTGT CAGGGCCTGC 720  
CAGACAGTGA GTCCTCTTTC ACATATACAC TAGATGAaAA GGTGgCCGAG TTAGTGGAGT 780  
TCCTGCTCCT CAAATACGAA GCAGAGGAGC CTGTAACAGA GGCAGAGATG CTGATGATTG 840  
TCATCAAGTA CAAAGATTAC TTTCTGTGA TACTCAAGAG AGCCCGTGAG TTCATGGAGC 900  
TTCTTTTTGG CTTGCCCCTG ATAGAAGTGG GCCCTGACCA CTTCTGTGTG TTTGCAAACA 960  
CAGTAGGCCT CACCGATGAG GGTAGTGATG ATGAGGGCAT GCCCGAGAAC AGCCTCCTGA 1020  
TTATTATTCT GAGTGTGATC TTCATAAAGG GCAACTGTGC CTCTGAGGAG GTCATCTGGG 1080  
AAGTGCTGAA TGCAGTAGGG GTATATGCTG GGAGGGAGCA CTTCGTCTAT GGGGAGCCTA 1140  
GGGAGCTCCT CACTAAAGTT TGGGTGCAGG GACATTACCT GGAGTATCGG GAGGTGCCCC 1200  
ACAGTTCTCC TCCATATTAT GAATTCCTGT GGGGTCCAAG AGCCCATTC AAGAGCATCA 1260  
AGAAGAAAGT ACTAGAGTTT TTAGCCAAGC TGAACAACAC TGTTCTAGT TCCTTTCCAT 1320  
CCTGGTACAA GGATGCTTTG AAAGATGTGG AAGAGAGAGT CCAGGCCACA ATTGATACCG 1380  
CAGATGATGC CACTGTCATG GCCAGTGAAA GCCTCAGTGT CATGTCCAGC AACGTCTCCT 1440  
TTTCTGAGTG AAGTCTAGGA TAGTTTCTTC CCCTTGTTGTT TGAACAGGGC AGTTTAGGTT 1500

1005103.030402







GACACTGAGG AGGGCTGGGG GGAACGCCCC ACCTCAGAGG GCAGATTCCC AGAGATTCCC 300  
ACCCTGCTCC TCAAGTATCA GCCCTCGTAG AGCTCCCCAG TCAGCTCAGG CGGGGTGGCA 360  
GCCATCTTAT TCCTGGGTGA GTGGCGTAGG GGAGGCGGAG GCCTTGGTCT GAGGGTCCCA 420  
TGGCAAGTCA GCACGGGGAG CTGCCTCTGG TTGGCAGAGG GAAGATTCCC AGGCCCTGCT 480  
GGGGATAAGA CTGAGGAGTC ACATGTGCAT CAGAACGGAC GTGAGGCTAC CCCGACTGCC 540  
CCCATGGTAG AGTGCTGGGA GGTGGCTGCC ACCGCCCTAC CTCCCACTGC TCTCAGGGAT 600  
GTGGCGGTTG CTCTGAGGTT TTGCCTTAGG CCAGCAGAGT GGTGGAGGCT CGGCCCTCTC 660  
TGAGAAGCCG TGAAGTTGCT AATTAAATTC TGAGGGGGCC ATGCAGTCCA GAACTATGAG 720  
GCTCTGGGAT TCTGGCCAGC CCCAGCTGTC AGCCCTAGCA GGCCCAAGAC CCTACTTGCA 780  
GTCTTTAGCC TGAGGGGCTC CCTCACTTCC TCTTGCAGGT GCTCCAGGAA CCAGGTGGTG 840  
ACGAACTGGG TGTGAGGCAC ACAGCCTAAA GTCAGCACAG CAGAGGAGGC CCAGGCAGTG 900  
CCAGGAGTCA AGGTGAGTGC ACACCCTGGC TGTGTACCAA GGGCCCTACC CCCAGAAACA 960  
GAGGAGACCC CACAGCACCC GGCCCTACCC ACCTATTGTC ACTCCTGGGG TCTCAGGCTC 1020  
TGCTTGCCAG CTGTGCCCTG AGGTGTGTTT CCACATCCTC CTACAGGTTC CCAGCAGACA 1080  
AACTCCCTAG GAAGACAGGA GACCTGTGAG GCCCTAGAGC ACCACCTTAA GAGAAGAAGA 1140  
GCTGTAAGGT GGCCTTTGTC AGAGCCATCA TGGGTGAGTT TCTCAGCTGA GGCCACTCAC 1200  
ACTGTCACTC TCTTCCACAG GCCTGTTGGA TCTCATCATC CATATCCCTG TTGATACGTT 1260  
TACCTGCTGC TCCTGAAGAA GTCGTATGC CTCCCGTTCC AGGCGTTCCA TTCCGCAACG 1320  
TTGACAACGA CTCCCCGACC TCAGTTGAGT TAGAAGACTG GGTAGATGCA CAGCATCCCA 1380  
CAGATGAGGA AGAGGAGGAA GCCTCCTCCG CCTCTTCCAC TTTGTACTTA GTATTTTCCC 1440  
CCTCTTCTTT CTCCACATCC TCTTCTCTGA TTCTTGGTGG TCCTGAGGAG GAGGAGGTGC 1500  
CCTCTGGTGT GATACCAAAT CTTACCGAGA GCATTCCCAG TAGTCCTCCA CAGGGTCCTC 1560  
CACAGGGTCC TTCCCAGAGT CCTCTGAGCT CCTGCTGCTC CTCTTTTTC TGGAGCTCAT 1620  
TCAGTGAGGA GTCCAGCAGC CAGAAAGGGG AGGATACAGG CACCTGTCAG GGCCTGCCAG 1680  
ACAGTGAGTC CTCTTTCACA TATACACTAG ATGAAAAGGT GGCCGAGTTA GTGGAGTTCC 1740  
TGCTCCTCAA ATACGAAGCA GAGGAGCCTG TAACAGAGGC AGAGATGCTG ATGATTGTCA 1800  
TCAAGTACAA AGATTACTTT CCTGTGATAC TCAAGAGAGC CCGTGAGTTC ATGGAGCTTC 1860  
TTTTTGGCCT TGCCCTGATA GAAGTGGGCC CTGACCACTT CTGTGTGTTT GCAAACACAG 1920  
TAGGCCTCAC CGATGAGGGT AGTGATGATG AGGGCATGCC CGAGAACAGC CTCCTGATTA 1980  
TTATTCTGAG TGTGATCTTC ATAAAGGGCA ACTGTGCCTC TGAGGAGGTC ATCTGGGAAG 2040

TGCTGAATGC AGTAGGGGTA TATGCTGGGA GGGAGCACTT CGTCTATGGG GAGCCTAGGG 2100  
 AGCTCCTCAC TAAAGTTTGG GTGCAGGGAC ATTACCTGGA GTATCGGGAG GTGCCCCACA 2160  
 GTTCTCCTCC ATATTATGAA TTCCTGTGGG GTCCAAGAGC CCATTCAGAA AGCATCAAGA 2220  
 AGAAAGTACT AGAGTTTTTA GCCAAGCTGA ACAACACTGT TCCTAGTTCC TTTCCATCCT 2280  
 GGTACAAGGA TGCTTTGAAA GATGTGGAAG AGAGAGTCCA GGCCACAATT GATACCGCAG 2340  
 ATGATGCCAC TGTCATGGCC AGTGAAAGCC TCAGTGTCAT GTCCAGCAAC GTCTCCTTTT 2400  
 CTGAGTGAAG TCTAGGATAG TTTCTTCCCC TTGTGTTTGA ACAGGGCAGT TTAGGTTCTA 2460  
 GGTAGTGGAG GGCCAGGTGG GGCTCGAGGA ACGTAGTGTT CTTTGCAATTT CTGTCCCATA 2520  
 TGGGTGATGT AGAGATTAC CTGTTTTTCA GTATTTTCTA AATGCTTTTC CTTTGAATAG 2580  
 CAGGTAGTTA GCTTCAGAGT GTTAATTTAT GAATATTAGT CGCACATGTA TTGCTCTTTA 2640  
 TCTGGTTTAA GAGTAACAGT TTGATATTTT GTTAAAAAAA TGGAAATACC TTCTCCCTTA 2700  
 TTTTGTGATC TGTAACAGGG TAGTGTGGTA TTGTAATAGG CATTTTTTTTT TTTTTTTACA 2760  
 ATGTGCAATA ACTCAGCAGT TAAATAGTGG AACAAAATTG AAGGGTGGTC AGTAGTTTCA 2820  
 TTTCTTGTC CTGCTTATTC TTTTGTCTT GAAAATTATA TATACCTGGC TTTGCTTAGC 2880  
 TTGTTGAAGA AAGTAGCAGA AATTAAATCT TAATAAAAGA AAAAAAAAAA AAAAAAAGG 2940

(2) INFORMATION FOR SEQ ID NO: 21:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1041 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: SINGLE  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

ATGCCTCTCT TTCCAAACCT TCCACGCCTC AGCTTTGAGG AAGACTTCCA GAACCCGAGT 60  
 GTGACAGAGG ACTTGGTAGA TGCACAGGAT TCCATAGATG AGGAGGAGGA GGATGCCTCC 120  
 TCCACTTCCT CTCCTCTTTT CCACTTTTTA TCCCCCTCCT CCTCTTCCTT GTCCTCATCC 180  
 TCACCCTTGT CCTCACCCTT ACCCTCTACT CTCATTCTGG GTGTTCCAGA AGATGAGGAT 240  
 ATGCCTGCTG CTGGGATGCC ACCTCTTCCC CAGAGTCCTG CTGAGATTCC TCCCCAGGGT 300  
 CCTCCCAAGA TCTCTCCCCA GGGTCCTCCG CAGAGTCCTC CCCAGAGTCC TCTAGACTCC 360  
 TGCTCATCCC CTCTTTTGTG GACCCGATTG GATGAGGAGT CCAGCAGTGA AGAGGAGGAT 420  
 ACAGCTACTT GGCATGCCTT GCCAGAAAGT GAATCCTTGC CCAGGTATGC CCTGGATGAA 480  
 AAGGTGGCTG AGTTGGTGCA GTTTCTTCTC CTCAAATATC AAACAAAAGA GCCTGTCACA 540

AAGGCAGAGA TGCTGACGAC TGTCAATCAAG AAGTATAAGG ACTATTTTCC CATGATCTTC 600  
 GGGAAAGCCC ATGAGTTCAT AGAGCTAATT TTTGGCATTG CCCTGACTGA TATGGACCCC 660  
 GACAACCACT CCTATTTCTT TGAAGACACA TTAGACCTCA CCTATGAGGG AAGCCTGATT 720  
 GATGACCAGG GCATGCCCAA GAACTGTCTC CTGATTCTTA TTCTCAGTAT GATCTTCATA 780  
 AAGGGCAGCT GTGTCCCCGA GGAGGTCATC TGGGAAGTGT TGAGTGCAAT AGGGGTGTGT 840  
 GCTGGGAGGG AGCACTTTAT ATATGGGGAT CCCAGAAAGC TGCTCACTAT ACATTGGGTG 900  
 CAGAGAAAGT ACCTGGAGTA CCGGGAGGTG CCCAACAGTG CTCCTCCACG TTATGAATTT 960  
 TTGTGGGGTC CAAGAGCCCA TTCAGAGGCC AGCAAGAGAA GTCTTAGAGT TTTTATCCAA 1020  
 GCTATCCAGT ATCATCCCTA G 1041

(2) INFORMATION FOR SEQ ID NO: 22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 346
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single-stranded
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

Met	Pro	Leu	Phe	Pro	Asn	Leu	Pro	Arg	Leu	Ser	Phe	Glu	Glu	Asp	Phe
				5					10					15	
Gln	Asn	Pro	Ser	Val	Thr	Glu	Asp	Leu	Val	Asp	Ala	Gln	Asp	Ser	Ile
			20					25					30		
Asp	Glu	Glu	Glu	Glu	Asp	Ala	Ser	Ser	Thr	Ser	Ser	Ser	Ser	Phe	His
		35					40					45			
Phe	Leu	Phe	Pro	Ser	Ser	Ser	Ser	Leu	Ser	Ser	Ser	Ser	Pro	Leu	Ser
	50					55				60					
Ser	Pro	Leu	Pro	Ser	Thr	Leu	Ile	Leu	Gly	Val	Pro	Glu	Asp	Glu	Asp
65					70					75					80
Met	Pro	Ala	Ala	Gly	Met	Pro	Pro	Leu	Pro	Gln	Ser	Pro	Pro	Glu	Ile
				85					90					95	
Pro	Pro	Gln	Gly	Pro	Pro	Lys	Ile	Ser	Pro	Gln	Gly	Pro	Pro	Gln	Ser
		100						105					110		
Pro	Pro	Gln	Ser	Pro	Leu	Asp	Ser	Cys	Ser	Ser	Pro	Leu	Leu	Trp	Thr
		115					120					125			
Arg	Leu	Asp	Glu	Glu	Ser	Ser	Ser	Glu	Glu	Glu	Asp	Thr	Ala	Thr	Trp
	130					135					140				
His	Ala	Leu	Pro	Glu	Ser	Glu	Ser	Leu	Pro	Arg	Tyr	Ala	Leu	Asp	Glu
145					150					155					160
Lys	Val	Ala	Glu	Leu	Val	Gln	Phe	Leu	Leu	Lys	Tyr	Gln	Thr	Lys	
			165					170					175		
Glu	Pro	Val	Thr	Lys	Ala	Glu	Met	Leu	Thr	Thr	Val	Ile	Lys	Lys	Tyr
		180					185						190		
Lys	Asp	Tyr	Phe	Pro	Met	Ile	Phe	Gly	Lys	Ala	His	Glu	Phe	Ile	Glu
	195						200					205			
Leu	Ile	Phe	Gly	Ile	Ala	Leu	Thr	Asp	Met	Asp	Pro	Asp	Asn	His	Ser
	210					215					220				
Tyr	Phe	Phe	Glu	Asp	Thr	Leu	Asp	Leu	Thr	Tyr	Glu	Gly	Ser	Leu	Ile
225					230					235				240	
Asp	Asp	Gln	Gly	Met	Pro	Lys	Asn	Cys	Leu	Leu	Ile	Leu	Ile	Leu	Ser

1005108 030102  
 201002 001500

**00000000000000000000000000000000**

(i) SEQUENCE CHARACTERISTICS:

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

(2) INFORMATION FOR SEQ ID NO: 24:

(i) SEQUENCE CHARACTERISTICS:

[illegible][illegible]

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1224 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single-stranded  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

ATGCCTCGGG GTCACAAGAG TAAGCTCCGT ACCTGTGAGA AACGCCAAGA GACCAATGGT 60  
CAGCCACAGG GTCTCACGGG TCCCCAGGCC ACTGCAGAGA AGCAGGAAGA GTCCCACTCT 120  
TCCTCATCCT CTTCTCGCGC TTGTCTGGGT GATTGTCGTA GGTCTTCTGA TGCCTCCATT 180

CCTCAGGAGT CTCAGGGAGT GTCACCCACT GGGTCTCCTG ATGCAGTTGT TTCATATTCA 240  
 AAATCCGATG TGGCTGCCAA CGGCCAAGAT GAGAAAAGTC CAAGCACCTC CCGTGATGCC 300  
 TCCGTTCTCTC AGGAGTCTCA GGGAGCTTCA CCCACTGGCT CTCCTGATGC AGGTGTTTCA 360  
 GGCTCAAAAT ATGATGTGGC TGCCAACGGC CAAGATGAGA AAAGTCCAAG CACTTCCCAT 420  
 GATGTCTCCG TTCCTCAGGA GTCTCAGGGA GCTTCACCCA CTGGCTCGCC TGATGCAGGT 480  
 GTTTCAGGCT CAAAATATGA TGTGGCTGCC GAGGGTGAAG ATGAGGAAAG TGTAAGCGCC 540  
 TCACAGAAAG CCATCATTTT TAAGCGCTTA AGCAAAGATG CTGTAAAGAA GAAGGCGTGC 600  
 ACGTTGGCGC AATTCCTGCA GAAGAAGTTT GAGAAGAAAG AGTCCATTTT GAAGGCAGAC 660  
 ATGCTGAAGT GTGTCCGAG AGAGTACAAG CCCTACTTCC CTCAGATCCT CAACAGAACC 720  
 TCCCAACATT TGGTGGTGGC CTTTGGCGTT GAATTGAAAG AAATGGATTG CAGCGGCGAG 780  
 TCCTACACCC TTGTCAGCAA GCTAGGCCTC CCCAGTGAAG GAATTCTGAG TGGTGATAAT 840  
 GCGCTGCCGA AGTCGGGTCT CCTGATGTCG CTCCTGGTTG TGATCTTCAT GAACGGCAAC 900  
 TGTGCCACTG AAGAGGAGGT CTGGGAGTTC CTGGGTCTGT TGGGGATATA TGATGGGATC 960  
 CTGCATTCAA TCTATGGGGA TGCTCGGAAG ATCATTACTG AAGATTTGGT GCAAGATAAG 1020  
 TACGTGGTTT ACCGGCAGGT GTGCAACAGT GATCCTCCAT GCTATGAGTT CCTGTGGGGT 1080  
 CCACGAGCCT ATGCTGAAAC CACCAAGATG AGAGTCCTGC GTGTTTTGGC CGACAGCAGT 1140  
 AACACCAGTC CCGGTTTATA CCCACATCTG TATGAAGACG CTTTGATAGA TGAGGTAGAG 1200  
 AGAGCATTGA GACTGAGAGC TTAA 1224

(2) INFORMATION FOR SEQ ID NO: 26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 407
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single-stranded
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

Met Pro Arg Gly His Lys Ser Lys Leu Arg Thr Cys Glu Lys Arg Gln  
 1 5 10 15

Glu Thr Asn Gly Gln Pro Gln Gly Leu Thr Gly Pro Gln Ala Thr Ala  
 20 25 30

Glu Lys Gln Glu Glu Ser His Ser Ser Ser Ser Ser Ser Arg Ala Cys  
 35 40 45

Leu Gly Asp Cys Arg Arg Ser Ser Asp Ala Ser Ile Pro Gln Glu Ser



50

55

60

Gln Gly Val Ser Pro Thr Gly Ser Pro Asp Ala Val Val Ser Tyr Ser  
65 70 75 80

Lys Ser Asp Val Ala Ala Asn Gly Gln Asp Glu Lys Ser Pro Ser Thr  
85 90 95

Ser Arg Asp Ala Ser Val Pro Gln Glu Ser Gln Gly Ala Ser Pro Thr  
100 105 110

Gly Ser Pro Asp Ala Gly Val Ser Gly Ser Lys Tyr Asp Val Ala Ala  
115 120 125

Asn Gly Gln Asp Glu Lys Ser Pro Ser Thr Ser His Asp Val Ser Val  
130 135 140

Pro Gln Glu Ser Gln Gly Ala Ser Pro Thr Gly Ser Pro Asp Ala Gly  
145 150 155 160

Val Ser Gly Ser Lys Tyr Asp Val Ala Ala Glu Gly Glu Asp Glu Glu  
165 170 175

Ser Val Ser Ala Ser Gln Lys Ala Ile Ile Phe Lys Arg Leu Ser Lys  
180 185 190

Asp Ala Val Lys Lys Lys Ala Cys Thr Leu Ala Gln Phe Leu Gln Lys  
195 200 205

Lys Phe Glu Lys Lys Glu Ser Ile Leu Lys Ala Asp Met Leu Lys Cys  
210 215 220

Val Arg Arg Glu Tyr Lys Pro Tyr Phe Pro Gln Ile Leu Asn Arg Thr  
225 230 235 240

Ser Gln His Leu Val Val Ala Phe Gly Val Glu Leu Lys Glu Met Asp  
245 250 255

Ser Ser Gly Glu Ser Tyr Thr Leu Val Ser Lys Leu Gly Leu Pro Ser  
260 265 270

Glu Gly Ile Leu Ser Gly Asp Asn Ala Leu Pro Lys Ser Gly Leu Leu  
275 280 285

Met Ser Leu Leu Val Val Ile Phe Met Asn Gly Asn Cys Ala Thr Glu

A vertical strip of 12 small, square, black-and-white photographs showing various stages of a plant's growth, from a seedling to a mature plant.

290

295

300

Glu Glu Val Trp Glu Phe Leu Gly Leu Leu Gly Ile Tyr Asp Gly Ile  
305 310 315 320

Leu His Ser Ile Tyr Gly Asp Ala Arg Lys Ile Ile Thr Glu Asp Leu  
325 330 335

Val Gln Asp Lys Tyr Val Val Tyr Arg Gln Val Cys Asn Ser Asp Pro  
340 345 350

Pro Cys Tyr Glu Phe Leu Trp Gly Pro Arg Ala Tyr Ala Glu Thr Thr  
355 360 365

Lys Met Arg Val Leu Arg Val Leu Ala Asp Ser Ser Asn Thr Ser Pro  
370 375 380

Gly Leu Tyr Pro His Leu Tyr Glu Asp Ala Leu Ile Asp Glu Val Glu  
385 390 395 400

Arg Ala Leu Arg Leu Arg Ala  
405

10085108 030402  
20100801 00150001